Checking if a Job was Killed by the OOM Killer

Category: Memory Usage on Pleiades

If a PBS job runs out of memory and is killed by the Out-Of-Memory (OOM) killer of the kernel, this event is likely (though not always) recorded in system log files. You can confirm this event by checking some of the messages recorded in system log files, and then increase your memory request in order to get your job running.

Follow the steps below to check whether your job has been killed by the OOM killer:

1. Find out when your job ran, what rack numbers were used by your job, and if the job exited with the Exit_status=137 from the tracejob output of your job. For example:

   ```
   pfe[1-12]% ssh pbspl1
   pbspl1% tracejob -n 3 140001
   ```

   where "3" indicates that you want to trace your job (PBS JOBID=140001), which ran within the past 3 days.

2. From the rack numbers (such as r2, r3, ...), you then grep messages that were recorded in the messages file stored in the leader node of those racks for your executable. For example, to look at messages for rack r2:

   ```
   pfe[1-12]% grep abc.exe /net/r2lead/var/log/messages
   Apr 21 00:32:50 r2i2n7 kernel: abc.exe invoked oom-killer:
   gfp_mask=0x201d2, order=0, oomkilladj=-17
   ```

3. Often, the Out-Of-Memory message doesn't make it into the messages file, but will be recorded in a consoles file named by each individual node. For example, to look for abc.exe invoking the OOM killer on node r2i2n7:

   ```
   pfe%  grep abc.exe /net/r2lead/var/log/consoles/r2i2n7
   abc.exe invoked oom-killer: gfp_mask=0x201d2, order=0, oomkilladj=0
   ```

   Note that these messages do not have a timestamp associated with them, so you will need to use an editor to view the file and look for the hourly time markers bracketing when the job ran out of memory. An hourly time marker looks like this:

   ```
   [-- MARK -- Thu Apr 21 00:00:00 2011]
   ```

It's also possible that a system process (such as, pbs_mom or ntpd) is listed as invoking the OOM killer, but it is nevertheless direct evidence that the node had run out of memory.

If you want to monitor the memory use of your job while it is running, you can use the tools listed in the article Memory Usage Overview.

In addition, NAS provides a script called *pbs_oom_check*. This script does the steps above and parses the /var/log/messages on all the nodes associated with pbs_jobid, looking for an instance of OOM killer. The script is available under /u/scicon/tools/bin and works best when run on the host pbspl1.

Article ID: 221
Last updated: 02 Aug, 2011
Computing at NAS -> Best Practices -> Memory Usage on Pleiades -> Checking if a Job was Killed by the OOM Killer
http://www.nas.nasa.gov/hecc/support/kb/entry/221/?ajax=1